**Rules for NR ligand screening (LXR, PPAR, RAR/RXR, FXR)**

1. Liver X receptor (LXR)

*Domain Check:*

* 4.7 <= VAIM <=7
* MW <= 750
* XLogP >= 2.0
* 5 <= TPSA <= 150

*Substructure Check:*

* “\*N[#6]1~[#6][#6](=O)N[#6]1=O” or “c1n[c,n]c2ccccc12” or “c1ccc2n[n,c]ccc2c1” or “a1aaaaa1S(=O)(=O)\*” or “a1aaaaa1CC(=O)O” or “Oa1aaaa1” or “O~Ca1aaaa1” or “Oa1aaaaa1” or “O~Ca1aaaaa1” or “\*C[#7](C\*)Ca1aaaa1” or “O[#6]~1~[#6]~[#6]~[#6]~[#6]~2~[#6]~3~[#6]~[#6]~[#6]~[#6]~[#6]3~[#6]~[#6]~[#6]12” or “O~C[#6]~1~[#6]~[#6]~[#6]~[#6]~2~[#6]~3~[#6]~[#6]~[#6]~[#6]~[#6]3~[#6]~[#6]~[#6]12” or “O[#6]~1~[#6]~[#6]~[#6]~2~[#6](~[#6]~[#6]~[#6]~3~[#6]~[#6]~[#6]~[#6]~[#6]~23)~[#6]1” or “O~C[#6]~1~[#6]~[#6]~[#6]~2~[#6](~[#6]~[#6]~[#6]~3~[#6]~[#6]~[#6]~[#6]~[#6]~23)~[#6]1” is present (fct. group or steroid)
* “[#6]~1~[#6]~[#6]~[#6]~[#6]~2~[#6]~3~[#6]~[#6]~[#6]~[#6]~[#6]3~[#6]~[#6]~[#6]12” or “c1ccccc1CC(F)(F)F” or “\*[#6](=O)c1ccc(O)cc1” or “a1aaaaa1~\*~\*~\*~\*~\*~\*~c1ccccc1” or “a1aaaaa1~\*~\*~\*~\*~\*~c1ccccc1” or “a1aaaaa1~\*~\*~\*~\*~c1ccccc1” or “a1aaaaa1~\*~\*~\*~c1ccccc1” or “a1aaaaa1~\*~\*~c1ccccc1” or “a1aaaaa1~\*~c1ccccc1” or “a1aaaa1~\*~\*~\*~\*~\*~\*~c1ccccc1” or “a1aaaa1~\*~\*~\*~\*~\*~c1ccccc1” or “a1aaaa1~\*~\*~\*~\*~c1ccccc1” or “a1aaaa1~\*~\*~\*~c1ccccc1” or “a1aaaa1~\*~\*~c1ccccc1” or “a1aaaa1~\*~c1ccccc1” (scaffold or steroid)

2. Peroxisome proliferator-activated receptor (PPAR)

*Domain Check:*

* 4.5 <= VAIM <=7
* MW <= 800
* 300 <= Ecc.Con.Index <=900
* 20 <= TPSA <= 300

*Substructure Check:*

* No “C~1~C~C~C2~C(~C1)~C~C~C1~C~C~C~C~C~2~1” is present (no steroids)
* “a(a)a~\*~\*~\*~\*~\*~\*~a(a)a” or “a(a)a~\*~\*~\*~\*~\*~a(a)a” or “a(a)a~\*~\*~\*~\*~a(a)a” or “a(a)a~\*~\*~\*~a(a)a” or “a(a)a~\*~\*~a(a)a” or “a(a)a~\*~a(a)a” is present (diaromatic scaffold) or “\*~[#6]~\*~[#6]~\*~[#6]~\*~[#6]~C(=O)[O,N]” or “[#6]~\*~[#6]~\*~[#6]~\*~[#6]~\*~C(=O)[O,N]” or “\*~[#6]~\*~[#6]~\*~[#6]~\*~[#6]a1aa([O,Cl,F,I,Br,N\*])aaa1” or “[#6]~\*~[#6]~\*~[#6]~\*~[#6]~\*~a1aa([O,Cl,F,I,Br,N\*])aaa1” or “\*~[#6]~\*~[#6]~\*~[#6]~\*~[#6]a1a([O,Cl,F,I,Br,N\*])aaaa1” or “[#6]~\*~[#6]~\*~[#6]~\*~[#6]~\*~a1a([O,Cl,F,I,Br,N\*])aaaa1” or “[#6][#6][#6][#6][#6][#6][#6][#6]NS(=O)(=O)N” (retinoid or fatty acid)
* “A[#8]a” or “a[#8]a” or “\*C(=O)O” or “c1nc2cncnc2n1” or “Oc1ccccc1” or “\*S(=O)(=O)N\*” or “\*S(=O)(=O)CN\*” or “\*S(=O)(=O)CCN\*” or “aa(a)[#7]” or “aa(a)~\*~[#7]” or “aa(a)~\*~\*~[#7]” or “aa(a)~\*~\*~\*~[#7]” or “aa(a)~\*~\*~\*~\*~[#7]” or “aa(a)~\*~\*~\*~\*~\*~[#7]” is present (fct. groups)

3. Retinoic acid/X receptor (RAR, RXR)

*Domain Check:*

* 5 <= VAIM <=7
* MV <= 630
* RB >= 3
* 300 <= Ecc.Con.Index <=1800
* XLogP >= 2.2 or alert “O~[#6]~[#6]~[#7]~[#6]” (amino-acid-type) is present but “\*[#6](~[#8])~[#6](~[#8])\*” (sugar-type) is not

*Substructure Check:*

* “\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~[#6](=O)~[#8]” or “\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~[#6](=O)~[#7]” or “\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~c1nnnn1” is present
* “\*1~\*([F,Cl,Br,I,C])~\*~\*~\*~\*~1” is present

3. Farnesoid X receptor (FXR)

*Domain Check:*

* MW <= 900
* 150 <= Ecc.Con.Index <=2400
* RB >= 2
* 15 <= TPSA <= 200

*Substructure Check:*

* “[#6]~1~[#6]~[#6]~[#6]2~[#6](~[#6]1)~[#6]~[#6]~[#6]1~[#6]~[#6](~[#8])~[#6]~[#6]~[#6]~2~1” or “[#6]~1~[#6]~[#6]~[#6]2~[#6](~[#6]1)~[#6]~[#6]~[#6]1~[#6]~[#6](~\*~[#8])~[#6]~[#6]~[#6]~2~1” or “[#6]~1~[#6]~[#6]~[#6]2~[#6](~[#6]1)~[#6]~[#6]~[#6]1~[#6]~[#6](~\*~\*~[#8])~[#6]~[#6]~[#6]~2~1” or “[#6]~1~[#6]~[#6]~[#6]2~[#6](~[#6]1)~[#6]~[#6]~[#6]1~[#6]~[#6](~\*~\*~\*~[#8])~[#6]~[#6]~[#6]~2~1”or “[#6]~1~[#6]~[#6]~[#6]2~[#6](~[#6]1)~[#6]~[#6]~[#6]1~[#6]~[#6](~\*~\*~\*~\*~[#8])~[#6]~[#6]~[#6]~2~1” or “[#6]~1~[#6]~[#6]~[#6]2~[#6](~[#6]1)~[#6]~[#6]~[#6]1~[#6]~[#6](~\*~\*~\*~\*~\*~[#8])~[#6]~[#6]~[#6]~2~1” or “[#8]~C(=O)~\*~\*~\*~\*~\*~\*~\*c1ccc([#8])cc1” or “[#8]~C(=O)~\*~\*~\*~\*~\*~\*~\*~\*c1ccc([#8])cc1” or “[#8]~C(=O)~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*c1ccc([#8])cc1” or “[#8]~C(=O)~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*c1ccc([#8])cc1” or “[#7]~C(=O)~\*~\*~\*~\*~\*~\*~\*c1ccc([#8])cc1” or “[#7]~C(=O)~\*~\*~\*~\*~\*~\*~\*~\*c1ccc([#8])cc1” or “[#7]~C(=O)~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*c1ccc([#8])cc1” or “[#7]~C(=O)~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*c1ccc([#8])cc1” or “[#8]=[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6](=[#8])[#8]” or “[#8]=[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6](=[#8])[#8]” or “[#8]=[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6](=[#8])[#8]” or “[#8]=[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6](=[#8])[#8]” or “[#8]=[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6]~[#6](=[#8])[#8]” is present (“bile salt” alert)
* “\*~\*~C(=O)Nc1ccccc1” or “\*c1nocc1CO\*” or “\*COCN1CC\*CC1” or “\*NC(=O)C(C)\*” or “NC(=O)c\*” or “\*cS(=O)(=O)Nc\*” or “OC1C~C~C(C(O)=O)C2(O)C~COC12” or “\*~[#7]C(=O)c1ccc(cc1)C(O)=O” or “\*~[#7]C(=O)c1ccc(cc1)\*C(O)=O” or “\*~[#7]C(=O)c1ccc(cc1)\*\*C(O)=O” or “\*~[#7]C(=O)c1ccc(cc1)\*\*\*C(O)=O” or “[#7]~\*~\*~\*~\*~c1ccc(cc1)C(O)=O” or “[#7]~\*~\*~\*~c1ccc(cc1)C(O)=O” or “[#7]~\*~\*~c1ccc(cc1)C(O)=O” or “[#7]~\*~c1ccc(cc1)C(O)=O” or “[#7]c1ccc(cc1)C(O)=O” or “[#7]~\*~\*~\*~\*~c1cc(ccc1)C(O)=O” or “[#7]~\*~\*~\*~c1cc(ccc1)C(O)=O” or “[#7]~\*~\*~c1cc(ccc1)C(O)=O” or “[#7]~\*~c1cc(ccc1)C(O)=O” or “[#7]c1cc(ccc1)C(O)=O” or “\*~C(=O)Nc1ccc(cc1)C(O)=O” or “\*~C(=O)Nc1ccc(cc1)\*C(O)=O” or “\*~C(=O)Nc1ccc(cc1)\*\*C(O)=O” or “\*~C(=O)Nc1ccc(cc1)\*\*\*C(O)=O” or “NCCc1c(CC(O)=O)[nH]c2ccccc12” or “\*C(=O)NC1CCCCC1” or “[#8]\*~\*~\*~\*~\*~\*c1nnnn1” or “[#8]\*~\*~\*~\*~\*~\*~\*c1nnnn1” or “[#8]\*~\*~\*~\*~\*~\*~\*~\*c1nnnn1” or “[#8]\*~\*~\*~\*~\*~\*~\*~\*~\*c1nnnn1” or “[#8]\*~\*~\*~\*~\*~\*~\*~\*~\*~\*c1nnnn1” or “[#8]\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*c1nnnn1” or “[#8]\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~\*~c1nnnn1” is present (assorted functional groups within scaffold)